

Wavelet 기반의 칼라와 질감 특징을 이용한 영상 검색

정 소 영*, 이 상 미**, 정 성 환*
* 창원대학교 전자계산학과 멀티미디어 연구실(MIPS)
** 한국전자통신연구원 정보유통연구실

Wavelet-based Image Retrieval Using Color and Texture Feature

So-young Jeong*, Sang-Mi Lee**, Sung-Hwan Jung*
* MIPS Lab., Dept. of computer Science, Changwon National University
** Electronics and Telecommunications Research Institute

요 약

영상검색을 위해 Wavelet 변환을 사용한 특징추출 접근방법은 영상들을 압축과 동시에 인덱스 할 수 있어서 영상 데이터베이스 저장과 관리의 복잡성이 상당히 감소될 수 있다. 본 연구는 각 영상의 Hue값에 대해 위치 정보와 주파수 정보를 가지는 Wavelet 변환의 성질을 이용하여 2단계 Wavelet 변환 후, 생성된 저대역 부밴드에서 칼라 특징을 추출하고 나머지 부밴드에서 질감 특징을 추출하여 영상 데이터베이스의 검색에 이용한다. 200개 영상을 사용하여 실험한 결과, 제안된 방법은 recall과 precision에서 약 97%, 81%의 검색 효율을 보였다.

I. 서 론

최근, 멀티미디어 기술과 더불어 영상 정보의 사용이 증가함에 따라 영상 데이터베이스로부터의 효과적인 영상 검색이 중요 관심사가 되고 있다. 기존의 데이터베이스들은 대부분이 문자 정보를 이용한다. 문자로 구성된 데이터 파일들은 알파벳 혹은 자음 순서로 정리되어 있고, 그 내용은 데이터베이스에서 저자와 키워드에 의해 인덱스된다. 인덱스 후 문자 매칭을 추출함으로써 원하는 정보를 검색할 수 있다. 이와 같은 접근 방식이 영상 데이터에 적용될 경우에는 몇가지 한계를 가진다. 먼저, 여러 가지 복잡한 특성을 가지는 영상의 경우 키워드에 의해 영상의 내용을 묘사하는 것이 어렵다. 두 번째, 대용량 데이터베이스에서 각 영상에 대한 주석은 많은 시간과 노력이 든다. 세 번째, 알려지지 않은 사람의 얼굴과 같은 샘플 영상에 기반한 질의의 경우가 있다. 이와 같은 문제의 해결을 위해 검색은 키워드보다 영상 특징에 기초되어야하므로 칼라, 모양, 질감 특징과 같은 low-level 영상 특징이 이용된다. 이

러한 접근방식을 내용기반 영상 검색이라고 한다[1].

내용 기반 접근 방법은 칼라, 모양, 질감 등을 통해 얻어진 특징 값을 사용하여 영상들간의 유사도를 계산하고 검색을 수행하므로 효과적인 정보 검색을 위해서는 영상 데이터를 나타내는 칼라, 모양, 질감 등 특징값의 추출 방법이 중요하다. 또한, 효율적인 영상 데이터 저장·관리와 특징 추출에 걸리는 시간을 최소화하기 위해서는 공간 영역보다는 변환 영역에서의 특징 추출에 대한 연구가 필요하다[2].

따라서, 본 논문에서는 변환 영역에서 영상의 내용 표현요소인 칼라와 질감 특징을 추출하여 데이터베이스내의 영상을 검색하는 방법을 제안한다. 칼라 특징 추출 방법에 있어서 디지털화된 영상들은 대부분 RGB 칼라 공간에서 명암도 값으로 표현되지만, 영상 검색에 있어서 HSI 칼라 공간이 더 효율적이기 때문에 본 논문에서는 hue(H), saturation(S), intensity(I)에 기초한 HSI 칼라 공간을 이

용한다[34].

서론에 이어, 제2장에서는 Wavelet 변환의 개요 및 내용기반 영상 검색을 위한 영상의 특징요소인 칼라와 질감 특징에 대해 서술한다. 제3장에서는 Wavelet 변환영역에서의 칼라와 질감 특징 추출 및 유사도 계산을 서술하고, 제4장에서는 실험결과를 기술한다. 마지막으로 제5장에서 결론을 맺고 앞으로의 연구과제에 대해서 살펴본다.

II. Wavelet 변환 및 영상 특징 요소

본 장에서는 Wavelet 변환의 개념 및 실제 영상에 대한 변환 수행 예를 보이고 각 영상의 특징 요소로서 칼라와 질감 특징에 대해 살펴본다.

2.1 Wavelet 변환 개요

에너지 압축, 다해상도 분석, 적응적인 공간-주파수장, 신호 decorrelation과 같은 몇가지 중요 특성을 가지는 Wavelet 변환은 데이터 압축을 위해 최근 많이 이용되고 있다.

Wavelet 변환 방법은, 영상의 변환 후 생성되는 부밴드를 분석함으로써 영상 정보를 얻을 수 있다. 이는 Wavelet이라고 불리는 기본 함수(basis function)를 이동(translating)하고 확장(dilating)함으로써 주파수 영역에 따른 다해상도를 갖게 된다[5].

Wavelet 변환의 일반적인 수식은 다음과 같이 정의된다.

$$WT_{f(a,b)} = \int_{-\infty}^{\infty} \psi\left(\frac{t-b}{a}\right) f(t) dt \quad (1)$$

여기서, $\psi(t)$ 가 근원 Wavelet이고, $\psi(t)$ 를 b 만큼 이동하고 a 만큼 확장한다.

그림 1은 크기가 $X \times Y$ 인 원 영상의 2단계 Wavelet 분할 형태를 보인 것이다.

Wavelet 변환을 하면 먼저 1단계 분할에서 하나의 저대역 부밴드 O_1 과 수평, 수직, 대각선의 방향 성분을 가지는 3개의 고대역 부밴드 W_1^H, W_1^V, W_1^D 가 생성된다. 2단계 분할에서, 저대역 부밴드 O_1 은 하나의 저대역 부밴드 O_2 와 3개의 고대역 부밴드 W_2^H, W_2^V, W_2^D 로 분해된다. 저대역 부밴드는 원 영상의 축소판으로 모든 Wavelet

부밴드 중에서 중요한 요소이다. 그리고 수평, 수직, 대각선 부밴드들은 각각 수평, 수직, 대각선 방향의 영상 에지 정보를 제공한다.

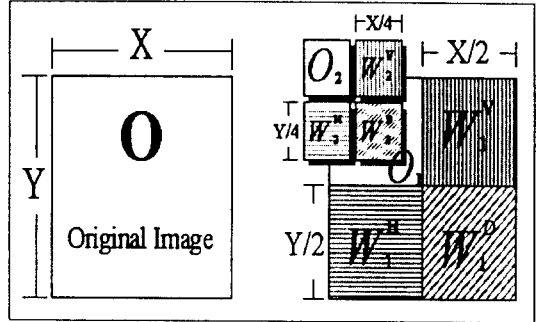


Fig. 1. 2-level Wavelet transform

그림 2는 실제 영상에 대하여 Wavelet 변환을 수행한 후의 분할 형태를 보인 예이다.

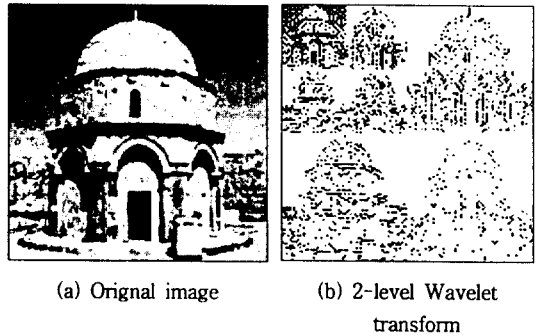


Fig. 2. An example of 2-level Wavelet decomposition

그림 2(a)의 128×128 크기의 원 영상을 2단계 Wavelet 변환을 하면, 그림 2(b)와 같이 64×64 크기의 세 개의 부밴드와 32×32 크기의 네 개의 부밴드가 생성된다. 그림 2(b)에서 왼쪽 상단의 저대역 부밴드는 원 영상의 축소형태로 변환 후에도 대부분의 원 영상의 정보를 유지한다. 나머지 부밴드들은 영상의 수평, 수직, 대각선 방향의 에지 성분을 나타내고 있다.

저대역 부밴드에서 생성된 Wavelet 변환 영상은 원 영상보다 크기는 작지만 원 영상이 가지는 대부분 정보를 유지하고 있기 때문에 Wavelet 변환을 함으로써 저장공간과 처리 시간 면에서 효율성이 높아 대용량 영상 데이터

베이스에서 유용하게 사용될 수 있다.

2.2 칼라 특징

칼라 특징은 시스템에서 관련된 영상과 관련되지 않은 영상을 구별하기 위한 효과적인 특징으로 연구되어져 왔다. 칼라를 이용한 영상간의 유사성은 각 화소가 가지는 칼라 값(color value)의 차이에 의해 판단되며 현재 대부분의 연구들은 칼라 값을 얻기 위하여 히스토그램 방법을 이용한다[6]. 히스토그램 방법은 알고리즘이 간단한 반면, 물체의 회전이나 크기 변화 등에 비교적 강한 편이다. 그러나 연관된 화소들 사이의 위치 정보가 부족하기 때문에 영상의 2차원적인 공간 정보가 상실될 수 있다는 단점이 있다. 칼라 정보를 이용한 대표적인 연구로는 칼라 히스토그램을 사용한 1991년 Swain과 Ballard 등의 연구가 있었다[7]. 칼라의 위치 정보를 고려하기 위하여 영상의 지역적 칼라 분포 특성에 대한 연구도 있었으며[8], 근래에는 Smith, Chang 등에 의하여 칼라 및 혼합 특성을 이용한 영상의 효율적인 질의·검색 방법에 대한 연구도 시도되고 있다[9].

2.3 질감 특징

질감은 물체 표면의 거침 정도나 밝기 변화 정도를 나타내는 일정한 특성을 이루며 고유한 패턴으로 분류할 수 있다. 질감을 이용한 분석 방법은 크게 영상 원형의 규칙적인 공간상의 배열을 다루는 구조적인 방법과 영상내 각 화소 사이의 상호 관계를 분석하는 통계적인 방법으로 나눌 수 있다[10]. 구조적인 방법은 영상의 구조적인 원형(primitive)과 그들의 위치 규칙을 정의함으로써 질감 정보를 표현한다. 하지만 영상내 원형의 구조가 크고, 일정한 규칙성을 가지는 영상만이 분석 가능하다는 단점을 가진다. 그리고 통계적인 방법은 질감 성질을 묘사하기 위한 단순한 접근법으로, 한 영상을 대표하는 통계 값을 얻기 위하여 일반적으로 화소들간의 공간적인 상호 관계를 분석한다[11].

III. Wavelet 기반의 영상 검색 시스템

본 장에서는 Wavelet 변환 후 변환영역의 저대역 부밴드에서 칼라 특징을 추출하고 나머지 부밴드에서 질감 특징을 추출하는 방법과 영상 검색을 위한 유사도 계산에 대해 살펴본다. 그림 3은 본 논문에서 제안한 변환 영역에서의 칼라와 질감 특징을 이용한 검색 시스템의 구성도이다.

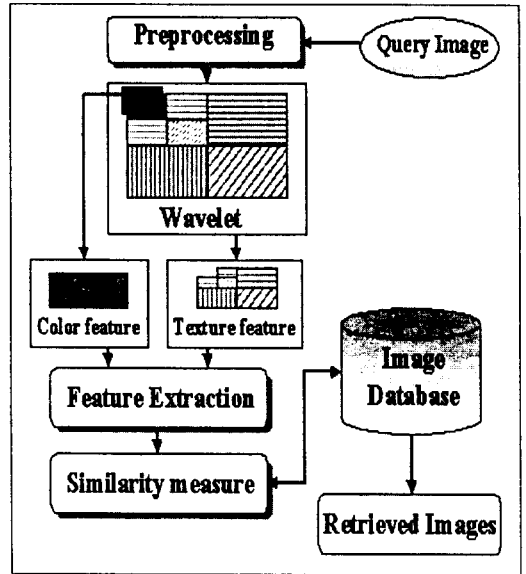


Fig. 3. Proposed image retrieval system

3.1 전처리 및 Wavelet 변환

영상의 특징을 얻기 위해, 전처리 단계에서는 해당 영상을 가로, 세로 128×128 크기로 정규화하고 RGB 칼라 공간에서 구성된 각 영상을 HSI 칼라공간으로 변환한 후 각 영상에 대한 Hue값을 Wavelet 변환의 입력 데이터로 한다. 각 영상의 효율적인 칼라 및 질감 특징을 구하기 위해 영상의 위치 정보와 주파수 정보를 가지는 다단계 Wavelet 변환을 수행한다.

3.2 특징 추출

각 영상에 대한 Hue값으로 Wavelet 변환을 한 후, 생성된 부밴드들 중에서 해당 영상의 칼라정보를 얻을 수 있는 저대역 부밴드에서는 칼라의 통계적인 방법을 이용하여 특징을 추출한다. 그리고 나머지 고대역 부밴드들에

는 질감의 통계적인 방법을 이용하여 특징을 추출한다.

3.2.1 칼라 특징 추출

저대역 부밴드에서 각 영상에 대하여 1차 모멘트, 2차 모멘트, 그리고 3차 모멘트값을 칼라 특징값으로 하고 각 특징 모멘트값으로서 영상의 칼라에 대한 평균, 표준편차, 그리고 왜도를 이용한다.

$$E_i = \frac{1}{A \cdot B} \sum_{j=1}^i P_{ij}, \quad S_i = \left(\frac{1}{A \cdot B} \sum_{j=1}^i (P_{ij} - E_i)^2 \right)^{\frac{1}{2}}, \quad (2)$$

$$\sigma_i = \left(\frac{1}{A \cdot B} \sum_{j=1}^i (P_{ij} - E_i)^3 \right)^{\frac{1}{3}}$$

여기서, E_i 는 평균, S_i 는 표준 편차, σ_i 는 왜도 그리고 P_{ij} 는 입력영상 $A \times B$ 의 j 번째 채널에 대한 i 번째 화소 값을 나타낸다.

3.2.2 질감 특징 추출

저대역 부밴드를 제외하고 나머지 부밴드들에서는 각 영상에 대하여 Wavelet 변환 계수의 Energy를 질감 특징값으로 얻는다[12]. 다음 식(3)은 질감 특징값을 얻기 위한 식이다.

$$E_k = \frac{1}{m \times n} \sum_i \sum_j W_k^2 [i, j] \quad (3)$$

여기서, E_k 는 나머지 6개 부밴드의 특징값들인 각각의 Energy를 나타낸다. 그리고, $m \times n$ 은 부밴드 영상들의 각각의 크기를 나타낸다. W_k 는 Wavelet 변환 후 각 부밴드의 변환계수를 나타낸다.

3.3 유사도 계산

영상의 검색을 위해서는 간단하면서도 효율적인 유사성 척도가 정의되어야 하고, 검색 시스템에 대한 효율성을 분석하기 위해서는 성능 평가 척도가 필요하다.

일반적으로 영상간의 유사도를 얻기 위한 척도로 각 영상의 특징값들 사이의 거리(distance)를 구하여 비교한다. 본 논문에서는 유사성 척도로서 Euclidean 거리 척도 함수의 변형인 간단한 City-block 거리 척도 함수를 이용한다[13].

$$D(Q, I)_{city-block} = \sum_i \sum_j |f - f'| \quad (4)$$

여기서, Q 는 질의 영상, I 는 데이터베이스내의 영상을 나타낸다. 그리고 f 와 f' 는 각각 두 영상의 특징값들이다.

칼라 특징값으로 이용된 평균, 표준편차, 왜도는 세자리 숫자를 넘지 않는 반면 질감 특징값으로 이용된 에너지는 다섯자리 숫자가 넘기 때문에 정규화해서 City-block 거리 척도 함수에 적용한다.

IV. 실험 결과 및 고찰

본 연구는 IBM-PC/586 컴퓨터 시스템과 GNU C언어를 사용하여 제안한 방법을 구현하였다. 실험에 사용된 영상 데이터는 WWW의 Corel Draw Photo Album 사이트에 분류되어 있는 200점의 영상을 사용하였다. 200점의 영상 데이터는 경주용 자동차 영상 57점, 하늘과 구름 영상 54점, 여러 종류의 꽃 영상 47점 그리고 여러 비행기 영상 42점으로 구성되었다. 각 실험 영상과 사용자 질의 영상은 RGB 칼라 bmp 파일로 128×128 크기로 정규화된 256 칼라를 가진다.

내용기반 영상 데이터 검색의 효율성 분석을 위한 성능 평가 척도로서 일반적으로 식(5)와 같은 Recall과 Precision을 이용한다. Recall은 영상 데이터베이스내에서 질의와 관련된 영상 중 검색된 영상의 비율을 말한다. 그리고 Precision은 검색된 영상 중에서 질의와 관련된 영상의 비율을 나타낸다[13].

$$Recall = \frac{R_r}{T}, \quad Precision = \frac{R_r}{T_r} \quad (5)$$

여기서, R_r 는 검색된 항목 중에서 질의와 관련된 항목의 수를 나타낸다. 그리고 T 는 검색 대상 중에서 질의와 관련된 항목의 총수를 말하고, T_r 은 검색된 항목의 총수를 나타낸다. 실험상에 관련된 유사 영상은 유사치를 각 영상에 대한 평균 이상으로 하였다.

전처리를 거친 질의 영상을 2단계 Wavelet 변환을 하여 하나의 저대역 부밴드와 여섯 개의 나머지 부밴드로 나눈다. 먼저 저대역 부밴드에서 칼라 특징으로 식(2)에 의해 평균, 표준 편차, 왜도를 구하고, 나머지 6개의 부밴드에서 질감 특징으로 식(3)에 의해 에너지를 계산한다. 구해진 특징값을 정규화한 후 정의된 유사도 계산식으로

유사도를 계산한다.

제안한 방법에 대한 성능 평가를 위해 앞서 정의한 성능 평가 척도로 기존의 Tamura 방법[14]과 비교했다. 표 1은 비교한 결과를 나타낸 것이다. 데이터베이스내에서 질의 영상에 대해 평균 유사도 이상을 가지는 영상을 검색 대상으로 한다.

Table 1. The comparison of the proposed method and Tamura's method

Method Class	Tamura		Proposed	
	Recall	Precision	Recall	Precision
Class 1	0.73	0.59	0.98	0.85
Class 2	0.74	0.60	0.97	0.84
Class 3	0.68	0.37	0.96	0.75
Class 4	0.52	0.31	0.97	0.73
Average	0.67	0.47	0.97	0.79

표 1에 나타난 것과 같이 제안한 방법이 Recall과 Precision에 있어 더 우수함을 알 수 있다.

그림 4는 질의 영상에 대해, 제안된 방법과 Tamura 방법 사이의 영상 검색 결과의 예를 보인 것이다.

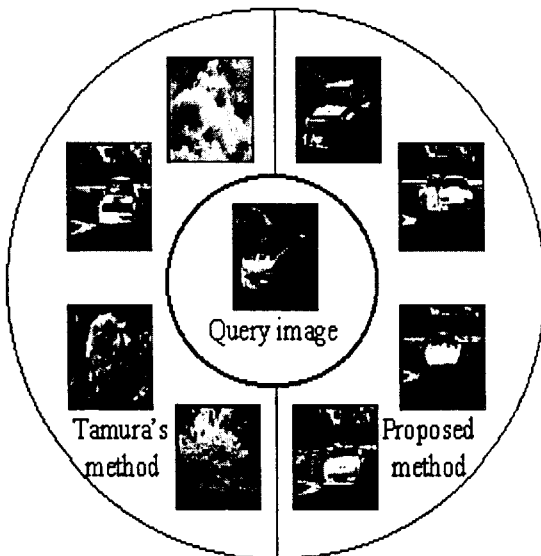


Fig. 4. An example of retrieval results.

V. 결론

본 논문에서는 내용 기반 영상 데이터 검색을 위하여 Wavelet 변환영역에서 칼라와 질감 특징을 이용하여 영상 데이터베이스를 검색하는 방법을 제안하였다.

먼저, 영상의 특징 추출을 위해 각 영상의 Hue에 대해 2단계 Wavelet 변환을 하여 하나의 저대역 부밴드와 나머지 6개의 부밴드를 얻는다. 원영상의 구조적 특징을 가장 많이 가지는 저대역 부밴드에서 칼라 특징으로 평균, 표준편차, 왜도를 구하고, 나머지 6개의 부밴드에서 질감 특징으로 에너지를 계산한다. 구해진 특징값을 정규화한 후 정의된 유사도 계산식으로 유사도를 계산하여 검색한다.

제안한 방법은 97%의 Recall과 81%의 Precision의 검색 효율을 보인다. 이는 Tamura의 방법과 비교하면 각각 45%, 71%의 성능 증가를 나타낸다.

본 연구는 Wavelet 변환을 사용함으로써 저장 공간을 줄여 멀티미디어의 대용량 영상 데이터베이스에서의 효율성을 기할 수 있고, 또한 칼라와 질감 특징을 동시에 사용함으로써 효율적인 검색을 할 수 있다.

앞으로의 연구는 다양한 영상에 대하여 검색 효율을 높일 수 있는 특징 추출법을 연구할 계획이다.

참고 문헌

- [1] Kai-Chieh Liang, X. Wan, and C.-C. Jay Kuo, "Indexing, retrieval and browsing of Wavelet compressed imagery data," SPIE Storage and Retrieval for Image and Video Databases V, Vol. 3022, pp.506-517, Feb. 1997.
- [2] M. K. Mandal and T. Aboulnasr, "Image Indexing Using Moments and Wavelets," IEEE Transactions on Consumer Electronics, Vol.42, No.3, pp.557-565, Aug. 1996.
- [3] Asha Vellaikal and C.-C. Jay Kuo, "Content-Based Image Retrieval Using Multiresolution Histogram Representation," SPIE Digital Image Storage and Archiving Systems, pp.312-323, Oct. 1995.
- [4] 김진숙, 김호성, "색과 공간정보를 이용한 영상 검색 기법," 한국정보과학회 '97 가을학술발표논문집, Vol.24, No.2, pp.483-486, 1997
- [5] M. K. Mandal, E. Chan, X. Wang and S.

- Panchanathan, "Multiresolution Motion Estimation Techniques for Video Compression," *Optical Engineering*, Vol.35, No.1, pp.128-136, January, 1988.
- [6] M. Stricker and M. Orengo, "Similarity of Color Image," *SPIE Vol.7*, No.1, pp.11-32, 1991.
- [7] M. J. Swain and D. H. Ballard, "Color Indexing," *International Journal of Computer Vision*. Vol.7, No.1, pp.11-32, 1991.
- [8] 배희정, 정성환, "지역 컬러 분포 특성을 이용한 컬러 영상 데이터 검색." *한국정보과학회 춘계학술발표논문집*, 제2권, 제1호, pp.306-309, 1995.
- [9] John R. Smith and Shih-Fu Chang, "Searching for Images and Videos on theWorld-Wide Web," *Center for Telecommunications Research Technical Report #459-96-25*, 1996.
- [10] 김희승, 영상인식-영상처리, 컴퓨터비전, 패턴인식, 신경망, 생능출판사, 1994.
- [11] Ramesh Jain, Rangachar Kasturi, and Brain G. Schunck, "Machine Vision," ISBN 0-07-032018-7, pp.234-248, 1995.
- [12] Ramesh Jain, S. N. Jayaram Murthy, and Peter L-J Chen, "Similarity Measures for Image Database," *SPIE Vol. 2420*, 1995
- [13] 김진아, 정성환, "Wavelet 기반의 질감 특징을 이용한 영상 검색," *한국정보과학회 '97 가을학술발표논문집*, Vol.24, No.2, pp.379-382, 1997
- [14] H. Tamura, S. Mori and T. Yamawaki, "Textures corresponding to visual perception," *IEEE Trans. Syst. Man Cybern. SMC-8(6)*, pp.460-473, 1978.